Driver Steering Behaviour Modelling Based on Neuromuscular Dynamics and Multi-Task Time-Series Transformer

Yang Xing[1] · Zhongxu Hu[2] · Xiaoyu Mo[2] · Peng Hang[3] · Shujing Li[4] · Yahui Liu[5] · Yifan Zhao[1] · Chen Lv[2]

Abstract

Driver steering intention prediction provides an augmented solution to the design of an onboard collaboration mechanism between human driver and intelligent vehicle. In this study, a multi-task sequential learning framework is developed to predict future steering torques and steering postures based on upper limb neuromuscular electromyography signals. The joint representation learning for driving postures and steering intention provides an in-depth understanding and accurate modelling of driving steering behaviours. Regarding different testing scenarios, two driving modes, namely, both-hand and single-right-hand modes, are studied. For each driving mode, three different driving postures are further evaluated. Next, a multi-task time-series transformer network (MTS-Trans) is developed to predict the future steering torques and driving postures based on the multi-variate sequential input and the self-attention mechanism. To evaluate the multi-task learning performance and information-sharing characteristics within the network, four distinct two-branch network architectures are evaluated. Empirical validation is conducted through a driving simulator-based experiment, encompassing 21 participants. The proposed model achieves accurate prediction results on future steering torque prediction as well as driving posture recognition for both two-hand and single-hand driving modes. These findings hold significant promise for the advancement of driver steering assistance systems, fostering mutual comprehension and synergy between human drivers and intelligent vehicles.

Keywords Driver steering behaviours · Neuromuscular dynamics · Multi-task learning · Sequential transformer · Intelligent vehicles

Abbreviations

ADAS	Advanced driver assistance system
ADV	Automated driving vehicles
BRMSE	Balanced root mean square error.
EEG	Electroencephalogram
EMG	Electromyography
FC	Fully connected
FFNN	Feedforward neural network
GRU	Gated recurrent unit
LOO	Leave one out
LSTM	Long-short term memory
MTS	Multi-task time-series
MTS-Trans	Multi-task time-series transformer
RMSE	Root mean square error
RP	Random prediction

✉ Yang Xing
yang.x@cranfield.ac.uk

1. School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield MK43 0AL, United Kingdom

2. School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore

3. College of Transportation Engineering, Tongji University, Shanghai 200092, China

4. College of Computer Science and Technology, Jilin University, Changchun 130012, China

5. School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

1 Introduction

Intelligent vehicles are showing tremendous potential in the improvement of traffic safety, efficiency, and energy-saving [1]. Although many achievements have been obtained in the past decade, a challenging question still needs to be answered for future intelligent vehicles: how we can efficiently devise collaboration and interaction mechanisms between human drivers and increasingly autonomous intelligent vehicles [2]? Before fully automated driving vehicles (ADVs) can be realized, human drivers still need to share control authority with these vehicles. Although ADVs can

efficiently manage many driving tasks, the partial ADVs, for example, the L3/L4 ADVs according to SAE J3016 [3], still expect the human driver to assume control in emergencies. In this situation, mutual understanding in terms of anticipating the human physiological and psychological states such as intention, attention, and behaviours will enable the two agents (driver and intelligent vehicles) to interact efficiently and safely with each other [4].

The lateral steering collaboration between the driver and intelligent vehicles or partial ADVs contributes to safer steering control under normal and critical situations [5]. Jointly understanding human steering patterns and predicting driver steering intention can benefit the advanced driver assistance system (ADAS) and shared steering control system [6]. Meanwhile, the risk assessment and hazard prediction system can be developed based on the driver's steering intention and steering quality analysis, thus preventing the human driver from making dangerous steering manoeuvres [7]. Besides, as the driver shares steering control authority with ADVs, it is essential to estimate the steering intention and quality and provide proper assistance and compensation before the driver makes improper steering manoeuvres due to the lack of sufficient situation awareness. Such a function could be especially important when the driver has to take over control from secondary tasks or due to emergent collision avoidance purposes [8]. Meanwhile, correlations between the steering postures and steering efficiency and quality have been analysed [9]. Accordingly, the recognition of specific driving postures can also help to analyse driving/steering behaviours, fatigue, and effectiveness and can also help to understand the steering patterns and steering qualities so that accurate assistance can be provided.

Considering these superiorities, in this study, a multi-task deep time-series modelling approach for sequential steering torque and driving posture prediction is proposed to jointly learn the representation and leverage the domain-specific information for the two steering modes (single-hand and both-hand). The predicted steering torque and steering postures can then optimize the design of advanced shared steering control algorithms. Notably, it is found that a more accurate steering torque prediction can be achieved if the model jointly considers the steering postures during both the learning and inference stages. To further quantify the impact of steering postures, this study considers two driving modes (single-hand and both-hand) and six driving postures (three for each driving mode). Specifically, for the both-hand driving mode, the 3-clock, 10–10-clock, and 12-clock driving postures are studied. While for the single-hand driving mode, the 3-clock, 130-clock, and 12-clock driving postures are investigated. Detailed information will be discussed in Sect. 2.

The main contributions of this study can be summarized as follows. First, a multi-task time-series model is devised for predicting driver steering intention and recognizing driving postures. This framework builds a connection between the driver's neuromuscular dynamics and future steering torque over a certain horizon. Notably, this study represents the first attempt to jointly model steering postures and steering torch predictions, with a further exploration of multi-task learning framework design. Second, the impact of different driving modes and driving postures are studied based on the proposed model to exploit the impact of different driving postures on steering intention prediction. Last, quantitative analysis and comparison for the MTS-Trans frameworks are proposed considering different driving modes, postures, and features.

This study is organized as follows. The related works are discussed in Sect. 2. Then the detailed introduction to the experimental design, simulation platforms, and simulation scenarios are presented in Sect. 3. In Sect. 4, the proposed MTS-Trans frameworks for continuous steering intention prediction and driving posture recognition are discussed. In Sect. 5, the experimental results are evaluated. Finally, conclusions are presented in Sect. 6.

2 Related Works

Many existing studies on driver steering intention focus on the modelling of tactical behaviours and discrete intentions such as lane change and turn manoeuvres using driver models, traffic context perception, and vehicle dynamics information [10, 11]. Vision-based approaches have been widely studied in the past for attention and intention predictions [12, 13]. For instance, in Ref. [14], the vision and road context information were adopted for the estimation of five common driving manoeuvres by using a long-short term memory (LSTM)-based approach. An ensemble time-series network was proposed for real-time driver intention prediction in a highway environment [15]. In Ref. [16], a Bernoulli heatmap approach was developed for a convolution neural network for driver's head pose estimation. In Ref. [17], the relationship between eye gaze estimation and steering performance was also studied. In general, driver intention can be efficiently recognized with features from the in-cabin vision system, vehicle states, and human physical and cognitive states (e.g., heart rate and EEG) [18]. Vision-based systems are normally used to capture driver behaviours such as head pose, driving behaviours, and emotion, primarily due to their easy implementation and low-cost advantages [19]. However, the prediction of continuous steering intention, like steering torque prediction, remains relatively unexplored and cannot be efficiently solved with vision-based methods due to several reasons. For example, low-cost driver monitoring solutions, like the vision-based approaches, have natural difficulties in accurate and continuous steering behaviour prediction due to many visual restrictions, such as occlusion, calibration, diverse driving postures, and habits.

While some studies have explored vision-based approaches for anticipating driving events, researchers have also developed steering intention classification systems based on alternative sensor systems. For example, a brain-machine interface for predicting human driver steering intentions using EEG was developed in Ref. [20]. However, EEG-based approaches can struggle to establish direct connections between brain dynamics and neuromuscular steering dynamics. Similarly, vehicle dynamic data from the CAN bus, including steering force, vehicle heading, and lateral acceleration, can be used for accurate steering intention recognition. For instance, an efficient curve speed model was developed for a driving assistance system considering various driving styles [21]. Driver lane change manoeuvre can be recognized using an explicit mathematical model of the steering behaviour based on vehicle dynamic states [22]. However, using vehicle dynamics-based information makes it challenging to make advance predictions for steering intention, as the steering intention can only be detected after the driver has initiated the steering manoeuvre. Therefore, a more efficient continuous steering torque prediction method based on the electromyography (EMG) signals and the upper limb neuromuscular dynamics is proposed in this study to explore the potential for joint estimation of steering intention and steering postures.

Driver upper limb neuromuscular dynamics and their impact on steering behaviours have been widely studied in the past for advanced driver steering assistants. In Ref. [23], it was found that the human-perceived steering force during driving can be different from the physical force, and the EMG signals can be used for the perceived force estimation. Co-contraction behaviours of upper limb muscles were identified through EMG signal analysis in Ref. [24]. Then, muscle learning behaviours based on the co-contraction scheme for the unusual steering tasks were further observed in Ref. [25]. To evaluate the impact of driving postures on the steering tasks, muscle activity and efficient steering evaluation based on the EMG signals were developed [26]. It was found that the driver usually performs push steering manoeuvres when steering clockwise. Besides, muscle alternation and co-contraction were found to be a clue for the estimation of muscle workload. Then, an approach for estimating driver steering efficiency was developed using the EMG signals [27]. A multi-regression method was developed to build an effective model for estimating steering force and steering quality. In Ref. [28], Pick and Cole found that the steering force is heavily determined by two major muscles in the upper limb region. Then, a steering torque estimation model based on the multiple regression analysis was developed. A similar continuous steering intent prediction model for both-hand driving modes was developed in Ref. [29]. However, neither of these studies explores joint modelling for continuous steering intention prediction and steering posture

recognition. Moreover, how different driving postures and modes can influence steering intention prediction and sequential neuromuscular dynamic modelling remains an open question. Thus, to enhance long-term temporal steering feature representation, a multi-task Transformer-based network is designed in this study due to the powerful time-series modelling performance of the Transformer network [30, 31]. By considering different driving modes and driving postures, the proposed model offers insight into distinct steering patterns with different driving postures.

3 Experimental Design

In this section, the experimental setup for the simulator-based steering intention prediction system is presented, along with details on signal collection and EMG data processing.

3.1 Experiment Platform and Scenarios

An experiment testbed was developed using a six-degrees-of-freedom driving simulator. A detailed introduction of the platform can be found in Ref. [6], which can be used for a wide range of human-in-the-loop steering experiments. The CarSim system was used to develop the simulation environment. A DynPick WEF-6A1000 force sensor and TR-60TC torque angle sensor, which were implemented under the steering wheel of the driving simulator, were used to collect the real-time steering dynamics. The Nihon Kohden ZB-150H wireless sensors were used for the collection of EMG signals with a sampling frequency of 1000 Hz.

During the experiments, participants were required to drive the simulator with two different driving modes. For each driving mode, ten EMG electrodes were used to extract the upper limb neuromuscular dynamics, respectively. The detailed neuromuscular signals are summarized in Table 1. Twenty-one voluntary male participants were involved in the experiment. All the participants were deeply informed about the purpose and risks of this experiment and agreed to participate. The 21 participants were divided into three groups, namely, the skilled group, the average group, and the unskilled group, according to their reported driving experiences. The introduction of different driving experience improves the pattern diversity of the experiment data and help to avoid bias patterns in the analysis of driver neuromuscular dynamics.

All participants were instructed to control the steering wheel by following a constant sinusoidal input. The driving task assigned to them was a slalom manoeuvre conducted at a velocity of approximately 60 km/h, as depicted in Fig. 1. This task adheres to international standards for evaluating vehicle dynamics and driver control performance. The slalom task was chosen as it simulates scenarios in which

Table 1 Summary of detected neuromuscular signals for two driving modes

| Driving modes	Detected neuromuscular signals
Both-hand (left & right hands)	Anterior deltoid of the clavicular portion (PMA-C), deltoid anterior (DELT-A), deltoid posterior (DELT-P), teres major (TM), and long head of triceps brachii (TB-L)
Single-hand (right hand only)	PMA-C, DELT-A, DELT-P, TM, TB-L, deltoid middle (lateral) (DELT-M), triceps lateral head exterior (TB-LAT), biceps (BC), infraspinatus (INFT), pectoralis major (PM)
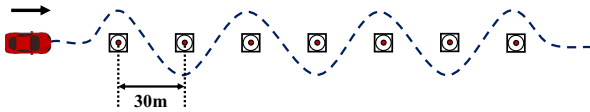
Fig.1 The Slalom driving task used in the simulation

drivers need to make lane changes or overtake other vehicles, common responses when taking over control from a partially automated driving vehicle. The critical aspect of intention prediction arises when a driver initially assumes control, making the slalom task an apt means of measuring driver steering behaviours in this study.

The participants were required to use both-hand driving mode in the first with hands on the 3-clock positions, as shown in Fig. 2(d). Then, they should perform a four to six-second steering experiment by following a constant period angle signal. The target sine-wave-like signal has about a 60-degree amplitude with a 0.25 Hz frequency to mimic real-world steering behaviour. The driving cycle will be repeated three to five times for each participant, respectively. In general, each participant will be measured with a 12 to 15 period of steering with around five seconds of steady holding phase after every three to five continuous steering. The main objective of this task is to simulate the real-world steering manoeuvre and to obtain the naturalistic neuromuscular dynamics from the participants. The data that support all the findings of this study are available from the corresponding authors upon reasonable research-purpose request.

3.2 EMG Signal Processing and Time-Delay Analysis

Before the commencement of the experiment, careful placement of the EMG electrodes on each participant is ensured. Subsequently, participants can take at least three minutes to prepare themselves on the testbed and sit in a preferred position to make sure they are relaxed. When the experiment starts, the EMG signals are collected and checked. The baseline noise is recorded and filtered for each channel. The steering torque for each participant is ranged within -5 to 5 N·m, and the EMG signals are all ranged within -5 to 5 mV. The sampling frequency is 1000 Hz and a low-pass
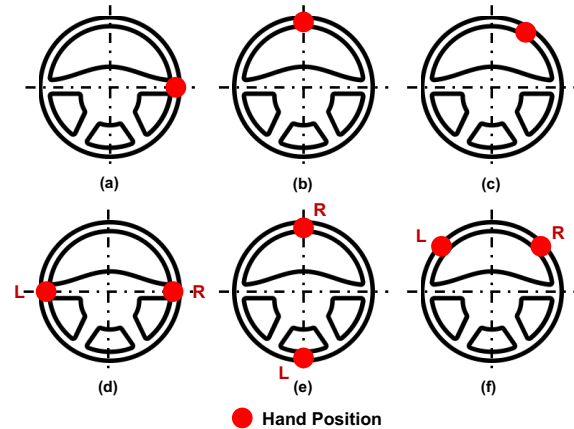
Fig. 2 Six driving postures for the two different driving modes. The red dot shows the hand grasp position. **a–c** single-hand driving mode with 3-clock, 12-clock, and 130-clock driving positions. **d–f** both-hand driving mode with 3-clock, 12-clock, and 10–10-clock driving posture, respectively

filter with 100 Hz pass-band frequency is then applied to the recorded EMG signals to filter the high-frequency outliers and interference. The time-series EMG data is divided into sequences with a fixed length of 200, which means a 200 ms prediction horizon without overlap is used as the sliding window for sequential data generation. The historical EMG and the steering torque signals will be used as the model input, and the future sequential steering torque is the target output of the model. All the signals are normalized within -1 to 1 before feeding into the model to avoid the adverse impact of a different unit.

The input (X) and target outputs $Y \epsilon [Y_t, Y_c]$ of the model can be denoted as:

$$X = \left[E_1, E_2, \ldots, E_9, E_{10}, Y_h\right] \tag{1}$$

$$E_i = [e_t, \ldots e_{t-200}] \tag{2}$$

$$Y_t = [st_{t+1}, \cdots, st_{t+200}] \tag{3}$$

where X is the model input at each time step, E_i is the ith EMG feature sequence that consists of the last 200 ms, Y_t is

predicted torque, which contains the future 200 ms torque information, Y_h is the historical steering torques in the past 200 ms, and st_{t+1} is the steering torque at time $t + 1$. Hence, the dimension of the model input sequence (X) is 11×200, and the dimension of the output sequence (Y_t) is 1×200. Here Y_c represents the predicted driving postures, which can be one of the three categories for the both-hand and single-hand driving modes, as shown in Fig. 2.

The initial analysis undertaken in this study involved measuring the time delay between the EMG signals and steering torque. The aim is to estimate the relative phase shift (time delay) between the two data streams by identifying the point of maximum correlation. The largest phase shift reflects where the two signals are best aligned. A statistical illustration of the time delay between the 21 participants and the two driving modes is shown in Table 2. The mean and standard deviation of the time delay for each driving posture are reported. As shown in Table 2, for both-hand driving modes, the 3-clock hand position can generate the largest phase shift compared to the other two postures. On the contrary, for the single-hand driving mode, the 3-clock driving posture leads to the lowest time delay among the three groups, while the 12-clock steering posture generates significantly larger time delays (-292 ± 160 ms). Based on Table 2, the average time delay for the both-hand driving mode, considering the three postures, is about -197 ms, and the average time delay for the single-hand driving mode is -206 ms. It shows that the mean time delay between the steering torque and the EMG signals is about 200 ms for all three driving postures with both the both-hand and single-hand driving modes. Hence, the 200 ms prediction horizon is used for future steering torque prediction and evaluation in this study.

4 MTS-Trans Model Design

4.1 Multi-Task Learning for Driver Steering Behaviours Modelling

Multi-task learning (MTL) can improve model generalization ability and decrease the uncertainty and noise in the data by jointly optimizing multiple tasks. In accordance with
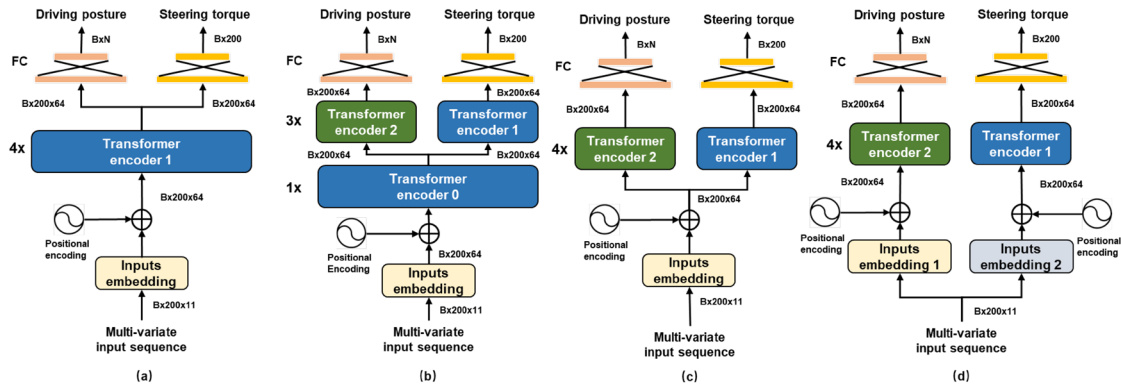
findings from a previous study [9], it is found that the steering postures are highly correlated with steering performance. Therefore, to accurately predict future steering torques and steering behaviours, the model will jointly learn how to recognize the steering postures to contribute to the accurate prediction of continuous steering behaviours. However, it is still unclear how these two tasks can be learned together to avoid any negative transfer effects [32]. It is also unclear whether the two tasks need to be learned at different rates or whether one task will dominate the learning and result in poorer performance on the other task [33].

Hence, the design of an MTL-based steering behaviour model necessitates an investigation into how much information these two tasks should share and the extent to which flexibility should be allotted to each task. As demonstrated in previous studies [34], a hard-MLT network shares the common backbone networks for the different tasks with relatively simple prediction heads for each task. As shown in Ref. [35], soft-MLT networks can further improve network performance by allowing information sharing and individual backbones for each task. Therefore, in the soft-MLT networks, sub-networks for different tasks may have more flexibility to tune their backbones [36]. It is also interesting to exploit that for the steering behaviour modelling, what kind of MTL framework can be used to benefit the joint optimization of the two different tasks, and what kind of information sharing level is required for the two tasks. Consequently, four different MTS-Trans models are designed in this study by varying the flexibility level of each sub-network. In summary, the basic MTS-Trans shares a common backbone network and only uses several fully connected layers as the prediction head. Then, a second MTS-Trans slightly increases the flexibility of each task by adding individual transformer networks on top of a shared transformer encoder. The last two more flexible MTS-Trans models assign individual transformer encoders, either with or without shared input embedding layers.

4.2 MTS-Trans Architecture

In this study, four different MTS-Trans architectures are developed, as shown in Fig. 3. Based on Eq. (1), a multi-variate input sequence $X \in \mathbb{R}^{l \times d}$ of length l and dimension d is fed into the sequential transformer networks. In general, all four architectures have similar blocks for the input embedding, transformer encoders, and fully connected prediction heads.

Embedding layer: the original embedding layer for discrete tokens in Ref. [30] is replaced with a fully connected layer. This fully connected layer is only applied to the feature dimension d and transforms the original 11 dimensions into 64 dimensions to adapt to the transformer network. Following the embedding layer, a standard sinusoid positional encoding is applied to inject positional information into the input

Table 2 Time delay (ms) analysis for different driving modes and postures

Driving modes	Position	Time Delay
Both-Hand	3-clock	-124 ± 109
	10–10-clock	-173 ± 192
	12-clock	-204 ± 174
Single-Hand	3-clock	-135 ± 116
	130-clock	-193 ± 161
	12-clock	-292 ± 160

Fig. 3 Four developed multi-task time-series transformer networks, respectively. **a** A shared embedding layer and three shared transformer encoder layers for the two tasks. **b** A shared embedding layer, one shared transformer encoder layer, and two personalized transformer encoder layers for the two tasks. **c** Different embedding and encoder layers for two tasks. **d** Has similar architecture with **b** but without the shared transformer encoder layer.

sequences [30]. The sine and cosine of different frequencies are calculated as follows:

$$PE(pos, 2i) = \sin(pos/10000^{2i/d_{\text{model}}}) \tag{4}$$

$$PE(pos, 2i + 1) = \cos(pos/10000^{2i/d_{\text{model}}}) \tag{5}$$

where *pos* represents the position of the sequential inputs, *i* is the ith dimension, and d_{model} is the overall feature dimension (64 in this study) after the linear embedding layer.

Transformer encoder: the transformer network is used here in order to capture both the long-term and short-term dependencies in the time-series input based on the multi-head self-attention mechanism. The transformer network can avoid sequential modelling and prediction mechanisms as used in LSTM but treat the past sentence as a whole and assign self-attention along with the sequences. The transformer is briefly introduced here, and the detailed information refers to Ref. [30]. The multi-head self-attention mechanism in the encoder output $H \in \mathbb{R}^{l \times d}$ of the same size as X by attending over given l key-value pairs $K \in \mathbb{R}^{l \times d}, V \in \mathbb{R}^{l \times d}$:

$$H = Attention(Q, K, V) = softmax\left(\frac{QK^{\text{T}}}{\sqrt{m}}\right)V \tag{6}$$

where Q, K, V are query, key, and value matrices for the self-attention module, respectively. The scaled-dot product attention in Eq. (6) allows compressing the matrix V into smaller representative bedding for simplified inference in downstream neural network operations. The scaled-dot product attention scales the weights by $\sqrt{m}$ to reduce the variance of the weights, facilitating stable training. Then the standard multi-head attention mechanism is applied to concatenate

several self-attention heads. In this study, four transformer encoder layers are used in the four different architectures and each layer has eight self-attention heads.

Fully connected prediction heads: on top of the transformer encoder layers are the FC prediction heads. The FC prediction heads follow the following architectures:

$$O_{dp} = fc_{2\_dp}(Dropout(Relu(fc_{1\_dp}(H_{\text{trans}})))) \tag{7}$$

$$O_{st} = fc_{2\_st}(Dropout(Relu(fc_{1\_st}(H_{\text{trans}})))) \tag{8}$$

where $O_{dp} \in \mathbb{R}^{N \times 1}$ and $O_{st} \in \mathbb{R}^{l \times 1}$ are predicted output for driving postures and steering torque, N is the number of driving postures. The output state H_{trans} from the transformer, the layer is flattened into a one-dimension vector and passed into the two FC layers. The output size of fc_{1_dp} and fc_{1_st} is 1024 uniformly.

4.3 Loss Function

The optimization of the MTS-Trans model follows a multi-task learning framework. The two tasks are jointly optimized, and the model is trained in an end-to-end fashion. The overall training loss L_T is a combination of the two individual losses, where the Cross-Entropy loss Eq. (9) is used for the posture recognition and the mean squared error Eq. (10) is used for the steering torque prediction. To deal with the homoscedastic uncertainty during model learning, the weighted loss function is used [34]. The overall loss function for MTS-Trans is denoted as Eq. (11).

$$L_{\text{cla}} = -\sum_{n=1}^{N}\sum_{c=1}^{C} \log \frac{\exp(x_{n,yn})}{\sum_{c=1}^{C} \exp(x_{n,c})} \tag{9}$$

$$L_{\text{reg}} = \frac{1}{S} \sum_{i=0}^{S} (y_i - \widehat{y}_i)^2 \tag{10}$$

$$L_T = W_1 L_{\text{cla}} + W_2 L_{\text{reg}} + \sum_{i=1}^{nT} \log \sigma_i \tag{11}$$

where x is input, y is the target sequence, N is the minibatch dimensions, C is the number of classes, S is the number of samples, $\widehat{y}_i$ is predicted torque, and $W_i = \exp(-\log \sigma_i^2)$ is the trainable weight for each sub-loss term considering the homoscedastic uncertainty or the observation noise σ_i for the specific task (σ_i initialized to zero), L_{cla}, and L_{reg} are the loss of posture recognition and steering torque prediction, respectively.

4.4 Model Implementation

In this study, 2 a total of 4011 sequences are collected for the both-hand case from 20 participants. These sequences are divided into three categories, with 8749 samples for the 3-clock posture, 7614 samples for the 10–10-clock posture, and 7648 samples for the 12-clock posture. For the single-hand driving case, 25,886 data sequences were collected from 20 participants. Among these, 9207 sequences are generated from the 3-clock driving posture scenarios, 8766 sequences are from the 130-clock categories, and 7913 sequences are collected from the 12-clock scenarios. In the model training process, 80% of the sequence is used for training, while the remaining 20% is reserved for model validation. The leave-one-out (LOO) approach is used to randomly select one participant and is only used for real-time performance demonstration. The Adam optimizer is used for model optimization [37]. An initial learning rate (LR) of 0.001 is used, and a step LR schedule is implemented to decay the LR by a factor of 0.1 every 100 epochs. The batch size is 64 and the maximum epoch is 300. The model is developed using PyTorch and is trained on a single NVIDIA RTX A5000 GPU. In general, the training process takes around three hours to finish.

5 Evaluation and Discussion

5.1 Metrics and Baselines

To assess the accuracy of continuous torque predictions, two evaluation metrics, namely, the root-mean-square error (RMSE) and balanced RMSE (BRMSE), are employed. Additionally, for evaluating the performance of driving posture recognition, accuracy and $F1$ scores are utilized. Specifically, $RMSE$ is calculated as follows.

$$RMSE = \frac{1}{N} \sum_{j=1,2,\cdots N} \sqrt{\frac{1}{L} \sum_{i=1,2,\cdots L} \left(\left(\widehat{x}_{ji} - x_{ji} \right)^2 \right)} \tag{12}$$

where N is the number of sequences used for model testing, L is the length of each sequence, which is 200 according to the time delay analysis. $\widehat{x}_{ji}$ is the ith prediction in the sequence j, and x_{ji} is the ground truth value.

To avoid the adverse impact of the data unbalanced property on the model evaluation results, the *BRMSE* metric split the data into bins and calculated the mean *RMSE* for each bin, and the final average *BRMSE* is then calculated [38]. As the steering torque mostly ranged between -5 and 5 N·m, ten bins were selected with each bin size with a 1 N·m scale. The *BRMS* is represented as follows:

$$BRMSE_{d,k,j} = \frac{1}{N} \sum_i RMSE_i \tag{13}$$

$$BRMSE = \frac{1}{N_d} \sum BRMSE_{d,k,j} \tag{14}$$

where $BRMSE_{d,k,j}$ is the *BRMSE* for the jth bin in which $d = 1$ N·m is the bin size, $k = 10$ N·m is the maximum range of steering torque, $RMSE_i$ is the *RMSE* for ith sequence in the bin j, N is the number of sequences in jth bin, $N_d = 10$ is the overall number of bins, and *BRMSE* is the final calculated balanced *RMSE*. The accuracy and F1 score for the driving postures are defined as:

$$Acc = N_c / N_t \tag{15}$$

$$F1 = T_p / (T_p + (F_P + F_N)/2) \tag{16}$$

where N_c is the number of correct classifications, and N_t is the total number of samples. T_p, F_p, and F_N are the number of true positives, false positives, and false negatives, respectively.

To evaluate the model performance on the different tasks, several baseline methods are introduced below.

(1) **Random Prediction (RP)**. This baseline always predicts a zero torque value, which serves as a simple baseline. For posture recognition, it randomly selects one posture.

(2) **Feedforward Neural Network (FFNN)**. An FFNN model with a similar architecture to the prediction head of the MTS-Trans models.

(3) **LSTM Models (LSTM/Bi-LSTM)**. A two-layer LSTM model (with or without bi-directional connections) with 64 feature dimensions in a hidden state.

(4) **Gated Recurrent Unit (GRU) Models (GRU/Bi-GRU)**. Similarly, a two-layer GRU model (with or

without bi-directional connections) with 64 feature dimensions in the hidden state.

(5) **Single Task Transformer (STTrans).** These models have a structure similar to the MTS-Trans models as shown in Fig. 3, but they have only one prediction head at a time for single-task learning.

(6) **MTS-Trans.** These are the multi-task sequential transformer models with different fusion methods, as shown in Fig. 3.

5.2 Single-Hand Driving Results

The analysis of the single-hand driving mode is presented in Table 3. The results clearly demonstrate the effectiveness of the proposed MTS-Trans model in predicting steering torques and recognizing driving postures. Specifically, the MTS-Trans models achieved around 90% accuracy in posture recognition tasks. Meanwhile, accurate prediction of future steering torque is achieved within the range of 0.0678 N·m for *RMSE* with the MTS-Trans3 model. It is also shown that the third MTS-Trans model, which uses separate embedding layers and transformer encoder layers, generates the most accurate results compared to other methods.

Comparing multi-task learning with the single-task learning method, it becomes evident that multi-task learning can improve the overall model performance for both steering torque prediction and posture recognition. Specifically, when simultaneously learning the postures, the *RMSE* of the steering torque prediction decreased to 0.0678 N·m from 0.0794 N·m. Similarly, posture recognition accuracy increased from 87.74 to 90.16% when considering the steering torque task. This suggests that simultaneous learning of these two tasks and optimizing model parameters contributes to higher accuracy in both steering torque prediction and driving posture recognition.

The confusion matrix generated with the MTS-Trans4 model is shown in Fig. 4. An interesting pattern emerges where the 130-clock driving posture is more likely to be misclassified into the other two postures, and conversely, the other two postures are more likely to be misclassified as the 130-clock posture. Specifically, 117 samples from the 130-clock are misclassified into the 3-clock case, and 106 samples are misclassified into the 12-clock. For the 3-clock case, 156 samples are misclassified to the 130-clock samples, while only very few cases are misclassified into the 12-clock. Similar results can be found in the 12-clock case as well.

The results of classification accuracy and prediction error regarding different driving postures are depicted in Fig. 5. Specifically, Fig. 5(a), (b) evaluate the classification accuracy and prediction error for different models with the three different driving postures, respectively. A similar conclusion can be made that, among all the driving postures, the 130-clock is more likely to lead to worse performance than the other two postures on the two tasks (especially for the driving posture recognition). However, there is no significant difference between the postures of 3-clock and 12-clock for the single-hand driving mode.

The *BRMSE* results for the proposed MTS-Trans models are also evaluated. As shown in Fig. 6, the *BRMSE* for the four MTS-Trans models shows consistent results. Specifically, the minimum prediction errors are usually achieved between the 5th bin (prediction range from -1 to 0N·m) and 7th bin (1–2N·m), which shows the models are more easily to make much more accurate predictions and better capture the neuromuscular dynamics when expected steering torque or steering effort is low. Conversely, higher steering torque demands can lead to an increased prediction error for continuous steering intention. This finding provides evidence to support the notion

Table 3 Experiment results for the single-hand driving modes

Methods	Torque prediction			Posture recognition			
	RMSE (N·m)	5% *RMSE* (N·m)	*BRMSE* (N·m)	*Acc* (%)	Precision (%)	Recall (%)	*F*1 (%)
RP	2.4267	3.5882	2.4266	33.33	33.33	33.33	33.33
FFNN	0.1734	0.5404	0.1842	73.86	73.99	73.93	73.96
GRU	0.1022	0.3640	0.1157	72.88	72.58	73.11	72.64
LSTM	0.0832	0.1954	0.0896	84.79	84.86	84.83	84.84
BiGRU	0.0768	0.2619	0.0944	84.92	84.97	84.89	84.92
BiLSTM	0.0778	0.1047	0.0865	84.04	84.02	84.13	84.07
STTrans	0.0794	0.2723	0.0877	87.74	87.93	87.86	87.82
MTSTrans1	0.0892	0.1549	0.1102	89.73	89.69	89.78	89.73
MTSTrans2	0.0617	0.2310	0.0672	90.10	90.10	90.10	90.11
MTSTrans3	0.0737	0.02395	0.0793	89.40	89.41	89.37	89.37
MTSTrans4	**0.0564**	**0.2063**	**0.0622**	**90.10**	**90.20**	**90.14**	**90.16**

The bold values indicate the best results achieved by MTSTrans4 model

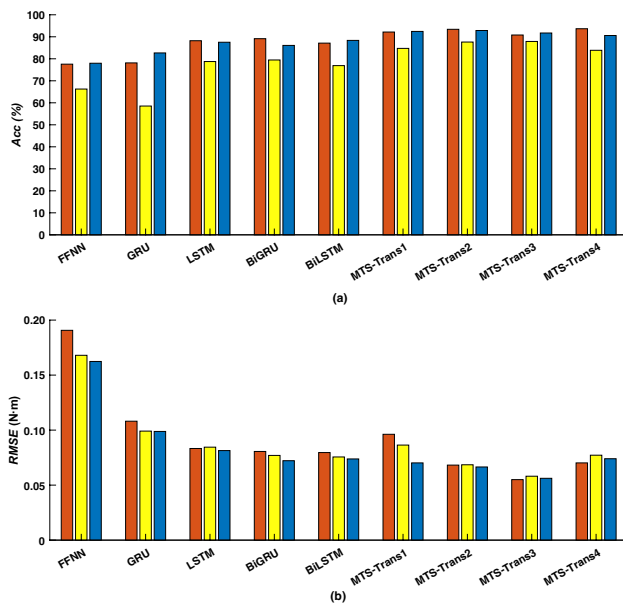**Fig. 4** Confusion matrix for the single-hand driving mode posture recognition with the MTS-Trans4 model

Fig. 5 Results comparison for driving postures recognition and steering torque prediction for different models and driving postures with single-hand mode. **a** Diving posture recognition accuracy for the three postures. **b** Prediction error (*RMSE*) for three different postures using different models

that upper limb neuromuscular dynamics exhibit relatively simple characteristics for a relatively light steering input. However, modelling these dynamics becomes more challenging when there is a need for heavy steering input. Hence, steering assistance from an intelligent vehicle can be expected when heavy steering demand is needed.

For the single-hand driving mode, several key observations can be made: The proposed MTS-Trans model achieves an approximate 90% recognition rate for the three
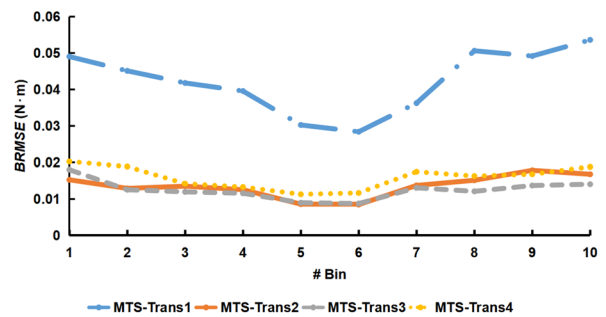

Fig. 6 Illustration of *BRMSE* prediction errors of four proposed MTS-Trans models with single-hand driving modes

postures while maintaining a prediction *RMSE* for steering torque at 0.0564 N·m; among the three postures, the 130-clock driving posture can lead to larger prediction errors and lower classification accuracy; the testing results for an unseen participant (employing the LOO approach) indicate that the proposed MTS-Trans model demonstrates strong generalization abilities for the single-hand driving modes, with most prediction errors occurring when higher steering torque is required.

5.3 Both-Hand Driving Results

In this section, the model developed for the both-hand driving mode is studied. Firstly, the steering torque prediction and posture recognition results based on different methods are presented in Table 4. It is shown that the MTS-Trans models show great advantages over the other methods in this case. Specifically, MTS-Trans3 and MTS-Trans 4 are the two most successful networks, showcasing exceptional joint optimization capabilities for both tasks. The *RMSE* values for the MTS-Trans3 and MTS-Trans4 models are 0.0399 N·m and 0.0401 N·m, respectively. The posture classification accuracy of the MTS-Trans3 model (97.47%) slightly outperforms that of the MTS-Trans4 model (97.22%). Compared to the single-task learning case (SSTrans), the proposed multi-task learning networks can increase overall performance on the two individual tasks. Besides, with the multi-task learning framework, the steering torque prediction performances are increased significantly (from 0.0901 to 0.0399 N·m) than the posture recognition.
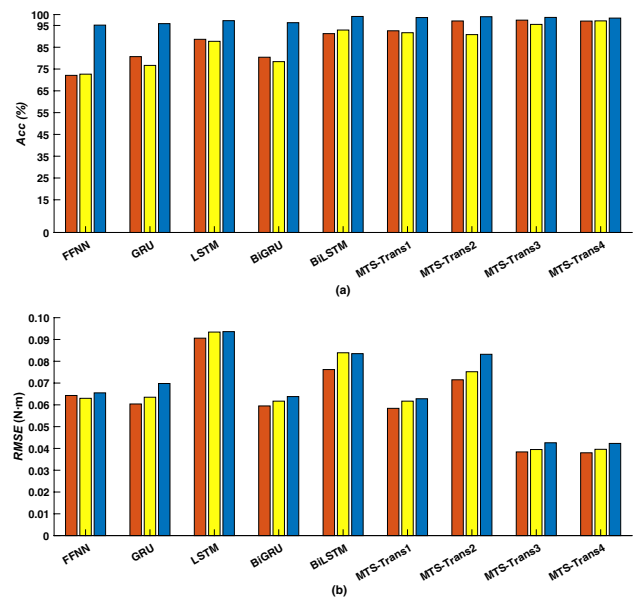
The confusion matrix generated with the MTS-Trans4 model is shown in Fig. 7. It is shown that for both-hand driving modes, 3-clock and 1010-clock postures are more likely to be misclassified and confused. However, the 12-clock posture is less likely to be misclassified into the 1010-clock but is more likely to be confused with the 3-clock posture (17 cases of 12-clock posture are misclassified into the 3-clock group while only 9 cases are misclassified into the 1010-clock group).

**Table 4** Experiment results for the both-hand driving mode

Methods	Torque prediction			Posture recognition			
	RMSE (N·m)	5% RMSE (N·m)	BRMSE (N·m)	Acc (%)	Precision (%)	Recall (%)	F1 (%)
RP	1.3576	3.5913	1.6851	33.33	33.33	33.33	33.33
FFNN	0.0643	0.1678	0.0744	79.60	79.50	79.97	79.69
GRU	0.0644	0.2251	0.0983	84.21	84.32	84.40	84.35
LSTM	0.0925	0.3216	0.1479	91.06	91.20	91.18	91.19
BiGRU	0.0616	0.2061	0.0884	84.81	84.84	85.03	84.92
BiLSTM	0.0810	0.2726	0.1238	94.28	94.30	94.43	94.36
STTrans	0.0901	0.2962	0.1416	95.52	95.65	95.53	95.58
MTSTrans1	0.0608	0.1981	0.0792	94.20	94.29	94.28	94.29
MTSTrans2	0.0764	0.2352	0.0919	95.68	95.93	95.62	95.72
MTSTrans3	0.0399	0.1336	0.0485	97.47	97.52	97.50	97.51
MTSTrans4	0.0401	0.1332	0.0471	97.22	97.29	97.21	97.25
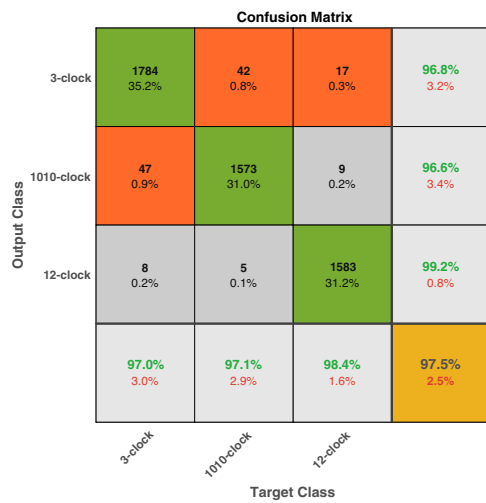
Fig. 7 Confusion matrix for both-hand driving mode posture recognition with the MTS-Trans3 model

Fig. 8 Results comparison for diving postures recognition and steering torque prediction for different models and driving postures with both-hand mode. **a** Diving posture recognition accuracy for three postures. **b** Prediction error (RMSE) for three postures using different models

The comparison of posture recognition accuracy and steering torque prediction RMSE regarding different driving postures and models are shown in Fig. 8. Contrary to the single-hand modes, it is interesting to observe that the 12-clock driving posture typically results in higher steering torque prediction errors than the other two postures, while also contributing to significantly higher posture recognition accuracy. This is notably different from the outcomes observed in the single-hand driving modes. Conversely, the 3-clock driving posture typically results in fewer steering torque prediction errors. Unlike the results seen in the single-hand driving mode, in the both-hand driving mode, the 3-clock driving postures appear to be better suited for posture steering torque prediction and understanding steering behaviour.

The BRMSE results for the proposed MTS-Trans models in both-hand driving modes are shown in Fig. 9. The BRMSE for the four MTS-Trans models exhibit similar results to those observed in the single-hand driving mode. Specifically, the minimum prediction errors are usually achieved between the 5th bin (within a prediction range of -1 to 0 N·m) and the 6th bin (0–1 N·m), while larger steering torque demand can lead to higher prediction error to the continuous steering intention.
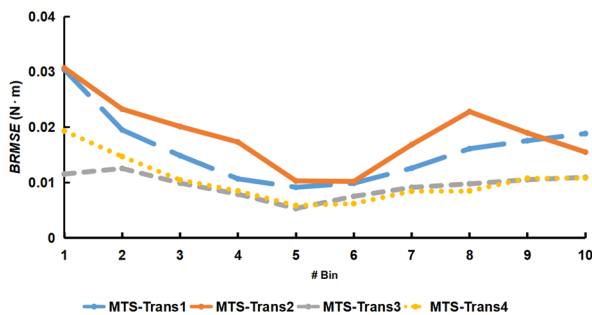
Fig. 9 Illustration of *BRMSE* prediction errors of four proposed MTS-Trans models with both-hand driving modes

The primary results for both-hand driving modes can be summarized as follows: The proposed MTS-Trans model can achieve a mean recognition accuracy of 97.47% for the three postures, with a prediction *RMSE* of 0.0399 N·m for steering torque. This results in more accurate steering torque prediction than the single-hand driving mode. The 3-clock driving posture results in fewer prediction errors than the other two methods, along with higher classification accuracy among the three postures. Based on the testing results for the unseen participant, it can be found that the proposed MTS-Trans model exhibits good generalization abilities in the single-hand driving mode, with most prediction errors occurring in specific areas.

5.4 Discussion

In this study, the sequential steering torque prediction and driving posture recognition model are developed based on a time-series transformer network. This model holds substantial significance as it enables an enhanced understanding between the driver and autonomous vehicle systems, primarily by predicting future steering torque, which serves as a proxy for the driver's steering intent. Moreover, based on the proposed multi-task steering behaviour model, the driving postures can be further recognized and their impact on steering intention prediction can be evaluated independently. This sequential steering torque prediction is instrumental not only in forecasting steering performance but also in facilitating the early detection of potential hazards or driver misbehaviour.

The advantages of the proposed system are summarized as follows: An EMG-based system connects the neuromuscular dynamics of the upper limb with steering torque and driving posture, which contributes to precise multi-task modelling for steering behaviours; the EMG-sensor-based approach is robust to the variation of situations and environments, such as the illumination and occlusion issues in vision-based systems; the proposed approach provides a theoretical analysis of upper limb neuromuscular dynamics

and how these dynamics can be applied and developed for driver steering intention prediction considering the different driving postures and hand positions on the steering wheel.

However, it's essential to acknowledge the potential concerns associated with the proposed EMG-based system, including intrusiveness, increased implementation time, and real-world applicability challenges. Addressing these concerns necessitates a forward-looking approach. This study has shown that by using the selected EMG signals, the number of EMG sensors can be reduced while similar steering torque prediction accuracy can be maintained. Future works can concentrate on more efficient feature engineering methods to select more efficient features for the EMG-based machine learning approach. Here are several prospective solutions:

(1) More challenging and critical steering manoeuvres can be designed to evaluate the robustness as well as the accuracy of the deep time-series model with upper limb EMG signals.
(2) The driving style has a critical impact on the steering behaviour as well [39]. It is also interesting to understand how to develop a personalized steering intention prediction model according to more complex driving style and experience aspects [40].
(3) With the development of wearable EMG sensors, such a system will become more suitable for real-world implementation [41].
(4) The EMG-based approach is more suitable for modelling individual neuromuscular dynamics and capturing the personalized characteristics of operating the vehicle or assistant mobility [42]. These characteristics will significantly benefit those special groups with a specific requirement or application domain, such as elder and disabled persons, and professional racing drivers [43].

6 Conclusions

This study presents a novel multi-task learning framework for modelling driving steering behaviour, employing a sequential transformer network. The proposed MTL-Trans model accurately estimates future steering torques and recognizes driving postures with upper limb neuromuscular dynamics. The investigation reveals that a prediction horizon of 200 ms effectively captures future steering torque based on historical observations. Remarkably, the MTS-Trans model achieves exceptional precision in steering torque prediction, yielding a low *RMSE* of 0.0564 N·m for single-hand driving and an even more impressive 0.0399 N·m for both-hand driving.

Furthermore, the proposed model exhibits a high degree of accuracy in identifying driving postures, attaining recognition rates of 97.47% and 90.10% for both-hand and single-hand driving modes, respectively. The research introduces a

quantitative assessment of the influence of driving postures on steering intent prediction, shedding light on how different postures impact steering behaviour modelling. Notably, this study underscores the capacity of driving postures to affect steering torque prediction accuracy, providing valuable insights into the intricate relationship between posture and steering behaviour modelling.

In conclusion, the MTL-Trans model's remarkable accuracy in predicting future steering torques and recognizing driving postures holds significant promise for the development of shared steering systems in intelligent vehicles. This framework enhances the potential for a more harmonious human-vehicle collaboration system, paving the way for safer and more efficient driving experiences.

Declarations

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest.

References

1. Rahman, M.H., Abdel-Aty, M., Wu, Y.: A multi-vehicle communication system to assess the safety and mobility of connected and automated vehicles. Transp. Res. Part C Emerg. Technol. **124**, 102887 (2021). https://doi.org/10.1016/j.trc.2020.102887

2. Li, W., Tan, R., Xing, Y., et al.: A multimodal psychological, physiological and behavioural dataset for human emotions in driving tasks. Sci. Data **9**, 481 (2022). https://doi.org/10.1038/s41597-022-01557-2

3. SAE, Taxonomy.: Definitions for terms related to driving automation systems for on-road motor vehicles. SAE Standard J 3016 (2016): 2016.

4. Hu, Z., Lou, S., Xing, Y., et al.: Review and perspectives on driver digital twin and its enabling technologies for intelligent vehicles. IEEE Trans. Intell. Veh. (2022). https://doi.org/10.1109/TIV.2022.3195635

5. Xing, Y., Lv, C., Cao, D., Hang, P.: Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving. Transp. Res. Part C Emerg. Technol. **128**, 103199 (2021). https://doi.org/10.1016/j.trc.2021.103199

6. Liu, Y., Liu, Q., Lv, C., Zheng, M., Ji, X.: A study on objective evaluation of vehicle steering comfort based on driver's electromyogram and movement trajectory. IEEE Trans. Human-Mach. Syst. **48**(1), 41–49 (2017). https://doi.org/10.1109/THMS.2017.2755469

7. Hang, P., Lv, C., Xing, Y., Huang, C., Hu, Z.: Human-like decision making for autonomous driving: a noncooperative game theoretic approach. IEEE Trans. Intell. Transp. Syst. Intell. Transp. Syst. **22**(4), 2076–2087 (2020). https://doi.org/10.1109/TITS.2020.3036984

8. Lv, C., Li, Y., Xing, Y., et al.: Human–machine collaboration for automated driving using an intelligent two-phase haptic interface. Adv. Intell. Syst. **3**(4), 2000229 (2021). https://doi.org/10.1002/aisy.202000229

9. Xing, Y., Lv, C., Zhao, Y., et al.: Pattern recognition and characterization of upper limb neuromuscular dynamics during driver-vehicle interactions. Iscience **23**(9), 101541 (2020). https://doi.org/10.1016/j.isci.2020.101541

10. Wang, W., Na, X., Cao, D., et al.: Decision-making in driver-automation shared control: a review and perspectives. IEEE/CAA J. Autom. Sin. **7**(5), 1289–1307 (2020). https://doi.org/10.1109/JAS.2020.1003294

11. Deng, Z., Hu, W., Yang, Y., et al.: Lane change decision-making with active interactions in dense highway traffic: a Bayesian game approach. In: 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), pp. 3290–3297. IEEE, (2022). https://doi.org/10.1109/ITSC55140.2022.9922333.

12. Xing, Y., Lv, C., Cao, D., Velenis, E.: Multi-scale driver behavior modeling based on deep spatial-temporal representation for intelligent vehicles. Transp. Res. Part C Emerg. Technol. **130**, 103288 (2021). https://doi.org/10.1016/j.trc.2021.103288

13. Hu, Z., Lv, C., Hang, P., Huang, C., et al.: Data-driven estimation of driver attention using calibration-free eye gaze and scene features. IEEE Trans. Ind. Electron. Ind. Electron. **69**(2), 1800–1808 (2021). https://doi.org/10.1109/TIE.2021.3057033

14. Jain, A., Koppula, H. S., Raghavan, B.,et al.: A car that knows before you do: Anticipating maneuvers via learning temporal driving models. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3182–3190 (2015)

15. Xing, Y., Lv, C., Wang, H., et al.: An ensemble deep learning approach for driver lane change intention inference. Transp. Res. Part C Emerg. Technol. **115**, 102615 (2020). https://doi.org/10.1016/j.trc.2020.102615

16. Hu, Z., Xing, Y., Lv, C., et al.: Deep convolutional neural network-based Bernoulli heatmap for head pose estimation. Neurocomputing **436**, 198–209 (2021). https://doi.org/10.1016/j.neucom.2021.01.048

17. Wang, Z., Zheng, R., Kaizuka, T., Nakano, K.: Relationship between gaze behavior and steering performance for driver–automation shared control: a driving simulator study. IEEE Trans. Intell. Veh. **4**(1), 154–166 (2018). https://doi.org/10.1109/TIV.2018.2886654

18. Li, W., Zeng, G., Zhang, J., et al.: Cogemonet: a cognitive-feature-augmented driver emotion recognition model for smart cockpit. IEEE Trans. Comput. Soc. Syst. **9**(3), 667–678 (2021). https://doi.org/10.1109/TCSS.2021.3127935

19. Hu, Z., Xing, Y., Gu, W., et al.: Driver anomaly quantification for intelligent vehicles: a contrastive learning approach with representation clustering. IEEE Trans. Intell. Veh. (2022). https://doi.org/10.1109/TIV.2022.3163458

20. Bi, L., Lu, Y., Fan, X., et al.: Queuing network modeling of driver EEG signals-based steering control. IEEE Trans. Neural Syst. Rehabil. Eng.Rehabil. Eng. **25**(8), 1117–1124 (2016). https://doi.org/10.1109/TNSRE.2016.2614003

21. Chu, D., Deng, Z., He, Y., et al.: Curve speed model for driver assistance based on driving style classification. IET Intel. Transport Syst. **11**(8), 501–510 (2017). https://doi.org/10.1049/iet-its.2016.0294

22. Zhao, Z., Zhou, L., Luo, Y., Li, K.: Emergency steering evasion assistance control based on driving behavior analysis. IEEE

Trans. Intell. Transp. Syst.Intell. Transp. Syst. **20**(2), 457–475 (2018). https://doi.org/10.1109/TITS.2018.2814687

23. Kishishita, Y., Takemura, K., Yamada, N., et al.: Prediction of perceived steering wheel operation force by muscle activity. IEEE Trans. Haptics **11**(4), 590–598 (2018). https://doi.org/10.1109/TOH.2018.2828425

24. Pick, A.J., Cole, D.J.: Neuromuscular dynamics in the driver–vehicle system. Veh. Syst. Dyn. Syst. Dyn. **44**(sup1), 624–631 (2006). https://doi.org/10.1080/00423110600882704

25. Pick, A.J., Cole, D.J.: Driver steering and muscle activity during a lane-change manoeuvre. Veh. Syst. Dyn. Syst. Dyn. **45**(9), 781–805 (2007). https://doi.org/10.1080/00423110601079276

26. Hayama, R., Liu, Y., Ji, X., et al.: Preliminary research on muscle activity in driver's steering maneuver for driver's assistance system evaluation. In: Proceedings of the FISITA 2012 World Automotive Congress: Volume 7: Vehicle Design and Testing (I), pp. 723–735. Springer, Berlin (2013). https://doi.org/10.1007/978-3-642-33835-9_66 .

27. Liu, Y., Ji, X., Hayama, R., Mizuno, T.: A novel estimating method for steering efficiency of the driver with electromyography signals. Chin. J. Mech. Eng. **27**(3), 460–467 (2014). https://doi.org/10.3901/CJME.2014.03.460

28. Pick, A.J., Cole, D.J.: Measurement of driver steering torque using electromyography. J. Dyn. Syst. Meas. Control (2006). https://doi.org/10.1115/1.2363198

29. Xing, Y., Lv, C., Liu, Y., et al.: Hybrid-learning-based driver steering intention prediction using neuromuscular dynamics. IEEE Trans. Ind. Electron. Ind. Electron. **69**(2), 1750–1761 (2021). https://doi.org/10.1109/TIE.2021.3059537

30. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. Advances in neural information processing systems, 30 (2017).

31. Zhou, H., Zhang, S., Peng, J., et al.: Informer: Beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 12, pp. 11106–11115 (2021). https://doi.org/10.1609/aaai.v35i12.17325

32. Kumar, A., Daume III, H.: Learning task grouping and overlap in multi-task learning. In: Proceedings of the 29th International Coference on Machine Learning (2012). https://doi.org/10.48550/arXiv.1206.6417

33. Standley, T., Zamir, A., Chen, D., et al: Which tasks should be learned together in multi-task learning? In: International Conference on Machine Learning, pp. 9120–9132. PMLR (2020)

34. Caruana, R.: Multitask learning: A knowledge-based source of inductive bias1. In: Proceedings of the Tenth International Conference on Machine Learning, pp. 41–48 (1993)

35. Yang, Y., Hospedales, T. M.: Trace norm regularised deep multi-task learning. In: 5th International Conference on Learning Representations (2016). https://doi.org/10.48550/arXiv.1606.04038

36. Kendall, A., Gal, Y., Cipolla, R.: Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7482–7491 (2018). https://doi.org/10.1109/CVPR.2018.00781

37. Kingma, D. P., Ba, J. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980 (2014)

38. Roth, M., Gavrila, D. M.: Dd-pose-a large-scale driver head pose benchmark. In: 2019 IEEE Intelligent Vehicles Symposium, pp. 927–934. IEEE (2019). https://doi.org/10.1109/IVS.2019.8814103

39. Deng, Z., Chu, D., Wu, C., et al.: A probabilistic model for driving-style-recognition-enabled driver steering behaviors. IEEE Trans. Syst. Man Cybernet. Syst. **52**(3), 1838–1851 (2020). https://doi.org/10.1109/TSMC.2020.3037229

40. Deng, Z., Chu, D., Wu, C., et al.: Curve safe speed model considering driving style based on driver behaviour questionnaire.

Transport. Res. F Traffic Psychol. Behav. **65**, 536–547 (2019). https://doi.org/10.1016/j.trf.2018.02.007

41. Hanbing, W., Yanhong, W., Xing, C., et al.: Human-vehicle dynamic model with driver's neuromuscular characteristic for shared control of autonomous vehicle. In: Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering (2020). https://doi.org/10.1177/0954407020977

42. Sanchez-Comas, A., Synnes, K., Hallberg, J.: Hardware for recognition of human activities: a review of smart home and AAL related technologies. Sensors **20**(15), 4227 (2020). https://doi.org/10.3390/s20154227

43. Taborri, J., Keogh, J., Kos, A., Santuz, A., et al.: Sport biomechanics applications using inertial, force, and EMG sensors: a literature overview. Appl. Bionics Biomech. (2020). https://doi.org/10.1155/2020/2041549

Yang Xing received his Ph.D. degree from Cranfield University, UK. He is currently a Lecturer in Applied Artificial Intelligence at Cranfield University. He was a Research Associate at the University of Oxford and a Research Fellow at Nanyang Technological University. His research interests include human behaviour modelling, intelligent multi-agent collaboration, and intelligent vehicles.

Zhongxu Hu received a mechatronic Ph.D. degree from the Huazhong University of Science and Technology of China, in 2018. He was a senior engineer at Huawei. He is currently a research fellow at the School of Mechanical and Aerospace Engineering of Nanyang Technological University in Singapore. His current research interests include human–machine collaboration, human cyber-physical systems, human digital twins, and deep learning applied in intelligent vehicle and industrial scenarios.

Xiaoyu Mo received the Ph.D. degree from Nanyang Technological University, Singapore, in 2022, the M.E. degree from Huazhong University of Science and Technology, Wuhan, China, in 2017, the B.E. degree from Yangzhou University, Yangzhou, China, in 2015. He is now a research fellow at Nanyang Technological University. His research interests include trajectory prediction and planning for connected autonomous vehicles.

Peng Hang is a Research Professor at the Department of Traffic Engineering, Tongji University, Shanghai, China. He received his Ph.D. degree with the School of Automotive Studies, Tongji University, Shanghai, China, in 2019. From 2020 to 2022, he served as a Research Fellow with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. His research interests include vehicle dynamics and control, decision-making, motion planning and motion control for autonomous vehicles.

Shujing Li received the B.S., M.S., and Ph.D. degrees from the College of Computer Science and Technology, Jilin University, China, in 2008, 2011, and 2020, respectively. He is currently a post-doctoral researcher in the College of Mathematics at Jilin University. He worked as an Assistant Engineer with the Jiangsu Institute of Automation, Lianyungang, China, from 2011 to 2012, and then worked as an Engineer from 2013 to 2017. He is interested in topics related to WBAN and VANET, especially for MAC layer protocols, and cross-layer optimization.

Yahui Liu received his B.S. degree from Jilin University, China, in 2003 and Ph.D. degree from Beihang University, China, in 2009 respectively. He is currently an Associate Professor at State Key Laboratory of Automotive Safety and Energy, Tsinghua University, China. His research interests include driver-vehicle system dynamics, steering system, and actuator design of ADAS.

Yifan Zhao was born in Zhejiang, China. He received the Ph.D. degree in automatic control and system engineering from the University of Sheffield, Sheffield, U.K., in 2007. He is currently a Reader in data science at Cranfield University, Cranfield, U.K. His research interests include computer-vision-based process monitoring, super resolution, active thermography, and nonlinear system identification.

Chen Lv is currently an assistant professor at Nanyang Technology University, Singapore. He received a Ph.D. degree at the Department of Automotive Engineering, Tsinghua University, China in 2016. From 2014 to 2015, he was a joint Ph.D. researcher at EECS Dept., University of California, Berkeley. His research focuses on cyber-physical systems, hybrid systems, advanced vehicle control, and intelligence, where he has contributed over 40 papers and obtained 11 granted China patents.